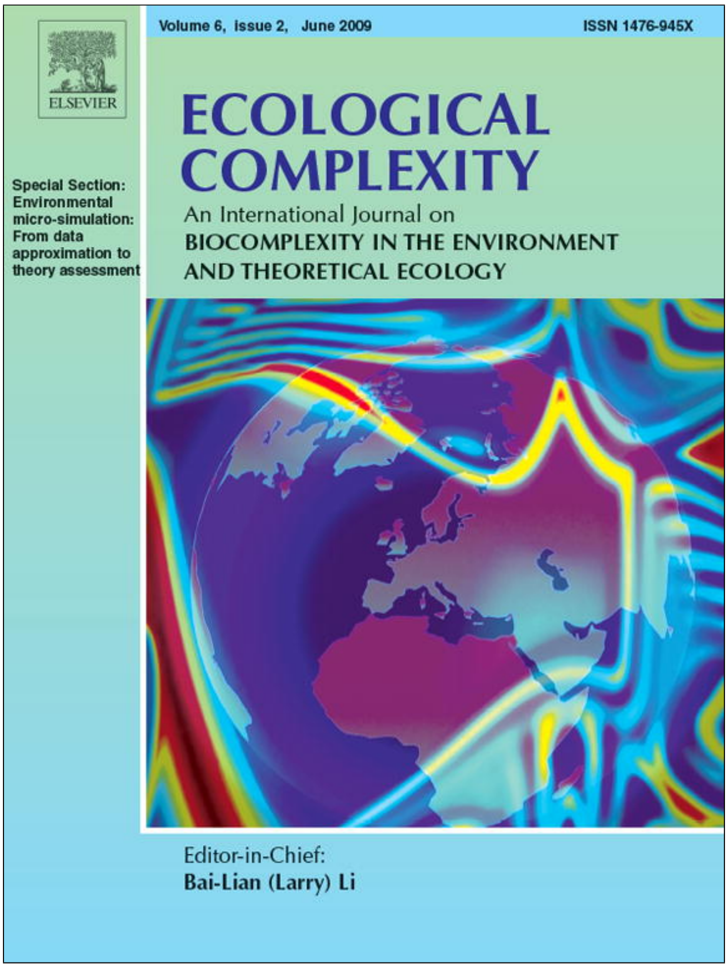
Volume 6, Issue 2, June 2009
ISSN 1476-945X
ELSEVIER
ECOLOGICAL COMPLEXITY
An International Journal on
BIOCOMPLEXITY IN THE ENVIRONMENT AND THEORETICAL ECOLOGY
Special Section: Environmental micro-simulation: From data approximation to theory assessment
Editor-in-Chief:
Bai-Lian (Larry) Li

Ecological Complexity 6 (2009) 77–79

Contents lists available at ScienceDirect

Ecological Complexity

journal homepage: www.elsevier.com/locate/ecocom

Scale and adequacy of environmental microsimulation

Tal Svoray [a,*], Itzhak Benenson [b]

[a] Department of Geography and Environmental Development, Ben-Gurion University of the Negev, Israel
[b] Department of Geography and Human Environment, Tel Aviv University, Israel

ARTICLE INFO

Article history:
Received 14 November 2008
Received in revised form 4 March 2009
Accepted 9 March 2009
Available online 19 April 2009

Keywords:
Environmental theory
Geographic information systems
Microsimulation
Remote sensing
Spatially explicit modeling
Spatial ecology

ABSTRACT

This special section of *Ecological Complexity* includes four papers that represent the use of microsimulation to study environmental problems. In the opening paper, we focus on the problem of environmental microsimulation (EMS) evaluation. We claim that the backbone of EMS lies in the developer's ability to capture the mechanisms that govern real-world dynamics and propose to substitute the real world by its virtual copy to investigate this ability. Namely, we shall employ EMS at a certain "likelihood" resolution and consider the unlimited data set generated by the model as a surrogate of reality. One can then sample this virtual world and investigate whether the methods of data analysis, model formulation, parameter estimation and model calibration, employed for the analysis of the real world, are sufficient to reveal *à priori* known mechanisms. A failure will manifest the inability of the researcher to capture real-world effects.

1. On environmental microsimulation

Microsimulation is a modeling technique that operates at the level of individual units, such as persons, households, vehicles, or firms (International Microsimulation Association—IMA website). In the last decade, environmental microsimulation (EMS), referring to the use of microsimulation to study environmental issues, has become a mainstream element of environmental studies and the use of these models is on the rise. To illustrate, Fig. 1 shows the increase in the number of papers that use the key words "environment + spatial + simulation + model,"[1] as cited in the ISI web-of-science database between 1996 and 2007. Especially strong growth is observed beginning from the year 2000 and we relate this to the intuitive fit of EMS to researchers' views of, on the one hand, environmental systems function and, on the other, improvements in computing capabilities, necessary for managing thousands of units that are simultaneously changing and, often, moving in space.

2. The validation delusion

Being spatially and temporally explicit, EMS rests on the availability of spatial data and an adequate representation of rules governing the behavior and interaction between the ecological and socio-economic systems in question. To represent a phenomenon, EMS development follows standard validation frameworks that are different from aggregate models, regarding the elementary model units. First, a formal description of the phenomenon is proposed. Each spatial unit of the system is characterized by several state variables and their change in time is represented by several parameterized dependencies (model rules) that describe the dynamics of the state variables in time. Second, the parameters of model rules are estimated, based on validation datasets. Third, the model is verified by reproducing the aggregate dynamics of the environmental pattern, usually of those that were previously employed for parameter estimation. Fourth, the model is used to investigate scenarios of future developments that always demand extrapolation of model dependencies.

This positivistic path enforces the developer's Occam's razor, so he or she always aims at minimizing the number of the unit state variables and rules parameters. However, to guarantee the sufficiency of such minimal parameterization, one needs experimental datasets that cover the entire domain of parameter variation. In spatially distributed systems considered by EMS, one should be sure that the model rules are valid over the whole modeled area. Inadequate description of the system dynamics over a part of it will result in representation adequate for the "measured islands" only. The risk is evident—one can never be sure of the importance of unmeasured parts for the overall dynamics of the region. The dynamics of these parts might dominate the entire system and result in uncontrolled deviation between models and real-world trajectories.

* *Corresponding author.* Tel.: +972 8 6472018; fax: +972 8 6472821.
E-mail addresses: tsvoray@bgu.ac.il (T. Svoray), bennya@post.tau.ac.il (I. Benenson).

[1] Being more generic, "spatial" has replaced "micro" in this search, as many authors do not use the latter as a keyword when referring to spatially explicit modeling.

1476-945X/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.ecocom.2009.03.011

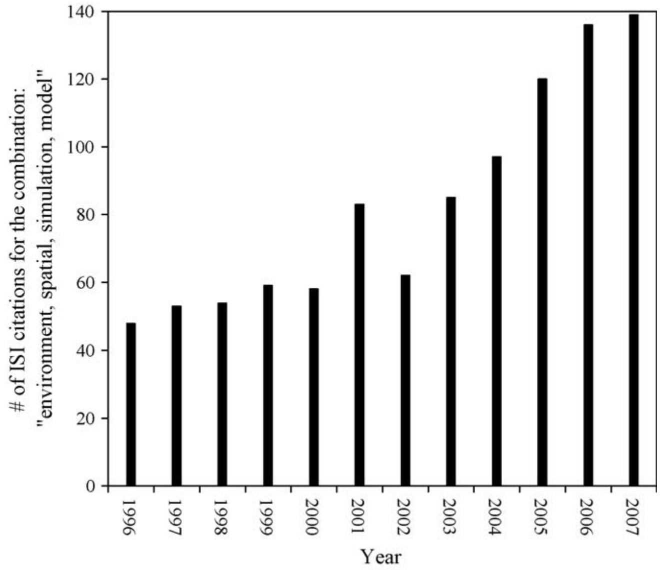

Fig. 1. A graph representing the frequency of research papers on environmental microsimulation, based on a search of the key words: Environmental, Spatial, Simulation, Model, in the ISI Web of Knowledge between the years 1996 and 2007.

However, it is hardly, if ever, possible to cover the entire area of model application with field observations. To take as an example, a case drawn from the field of rainfall, runoff and vegetation prediction, Fig. 2 presents two datasets obtained for the same rainfall catchment area. The first comprises field measurements of vegetation biomass, plotted against the contributing runoff area for several plots located along a semi-arid catchment in the center of Israel. The second presents the biomass values, as extracted from an IKONOS image, using NDVI, plotted against the corresponding runoff contributing area calculated from a DEM for every 10 m × 10 m unit of the IKONOS image. As can be seen, the locations of field measurements within the catchment area cover a wide, yet, insufficient spectrum of parameters. Some parts of the parameters' space are over-represented, while others are under-represented or not measured at all.

Benenson (2007) discussed similar problem regarding differences in spatial resolution of the real processes and their representation by EMS. Using the Game-of-Life example, he demonstrates the fact that a view of the system, at spatial or temporal resolutions coarser or finer than the "true" resolution of unit behavior, results in model rules that are completely different from the original rules of unit dynamics. He then raises the question: How one can know the "true" spatial resolution of the processes? If one cannot, does the fit between EMS models and data mean more than just an approximation of the observed data sets with the regression of "another kind"? And, we continue, how can one ascertain if our model rules really reflect those mechanisms that govern the environmental system?

Benenson's comments stress the relevance, for EMS, of the general problem of the unavoidably good fit of the multi-parametric models, raised by Douglas Lee for the case of non-spatial models in the mid 1970s (Lee, 1973, 1994). Each logically "reasonable" model that possesses a sufficiently high number of parameters can be fitted to the limited set of field measurements. A good model fit cannot, therefore, guarantee that the model rules reflect the underlying mechanisms of real-world dynamics.

That is, EMS models are only seemingly "better" than the standard comprehensive models of the 1970s and 1980s and the skeptical user could always regard them as just another tool of heuristic prediction of time series. The question we therefore pose is: Are we able to extract inherent mechanisms that govern a real-world system with calibration and verification procedures?

We believe that the rules that govern the real-world system are more robust in the face of changes in spatial and temporal resolution than the Game-of-Life and we propose to investigate our ability to reveal them, by studying the artificial world created by the model itself.

3. The alternative: M-world and its multi-scale views

To overcome the problem of the unavoidably good fit, we propose to substitute for the real world its virtual copy and then study our ability to recover its *à priori* known structure with the methods we employ when modeling the real world. This can be done by employing the outputs of a model M, employed at a highest possible resolution, as a surrogate for reality. In our field of investigation, this resolution is usually determined by the digital elevation model (DEM) of surface height and remote sensing images of the environmental conditions. The model output – the

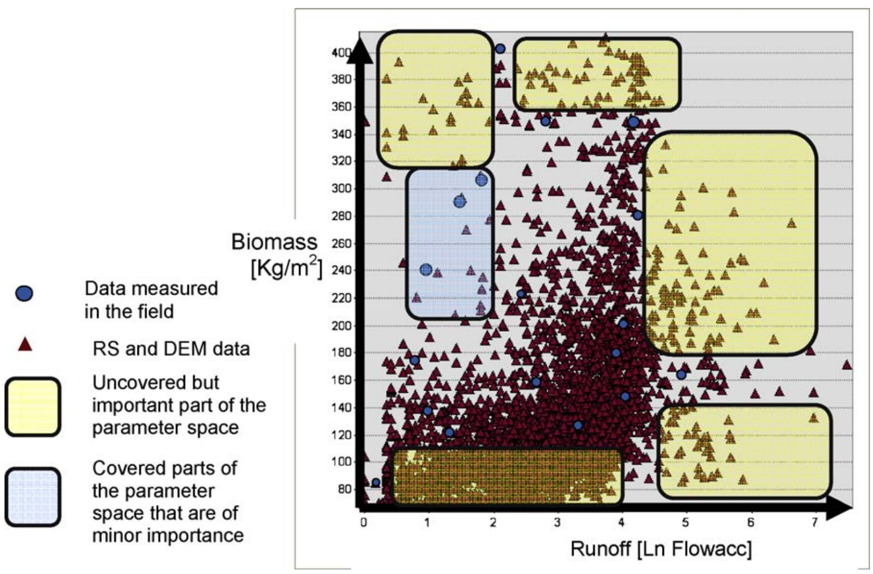

Fig. 2. A comparison between field data and remotely sensed data of the representation of the parameter space of a runoff contributing area versus vegetation production in a semi-arid site.

M-world – is externally indistinguishable from the (R)eal-world; however, the M-world database contains *all possible* trajectories characteristic of every combination of M-world parameters. The unlimited "experimental" data can be considered at the initial, or any lower resolution, to seek for the unambiguous methods of rules extraction and parameter estimation and to investigate the sensitivity of these methods to spatial and temporal resolution of the data.

4. M-world components

The M-world is based on two components. The *first* is a spatially and temporally explicit database that can be used for full multi-scale (multi-resolution) representations of the spatial phenomenon in question. With recent satellite data – at meter and even sub-meter resolutions – and with stereo-pair satellite and airborne data, the road is open for the use of land cover and topographical data as a standard component for constructing our EMS M-world. Additional components of the M-world database can be high-resolution vector GIS layers, accumulated during the last decade over vast areas in regards to land-uses and road networks.

The *second* component includes the set of *candidate* model mechanisms that represent our view of the real world. The model rules are usually based on the researcher's intuition and experience and, usually, the rules are ordered according to their likelihood. A minimal set of model rules is then chosen and the model parameters are estimated in a way that is also chosen by the researcher.

This is the stage on which we propose to focus. Let us consider the constructed model M at a highest possible spatio-temporal resolution and deploy the output of M as a surrogate for the reality. Is the spatial distribution of the experimental plots sufficient to recognize the rules that govern M? Then, let us generalize M's output and ask the researcher to estimate model rules and parameters again, based on the real world or modified pattern of the experimental plots. Are our theoretical views sufficient to construct the models rules that govern the entire M-world and not only the limited set of observations we have at hand? And, finally, do the extracted rules generate the correct prediction, when the model parameters are beyond the intervals available in observations? We do believe that the M-world studies may serve as litmus paper for the methods of model design as applied in EMS studies.

5. The papers

This special section focuses on simulating real-world phenomena with the help of GIS data. The papers consider dynamics of cities, epidemics, agricultural economics, and conservation of avian species, thus providing and verifying models of different environmental phenomena. Al-Ahmadi et al. (2009) apply urban cellular automata and fuzzy set theory to capture the uncertainty associated with the automata transition rules. The authors compare two methods of calibration: genetic algorithms and simulated annealing. Different patterns of urban development are studied and nine scenarios are devised to capture the effect of different development factors and their interactions.

The paper of Laperrière et al. (2009) investigates individual-based models of flea and rat behavior, aimed at description of plague transmission. The model enables a description of the disease dynamics and transmission at a level of individual organisms, in order to represent the collective dynamics of the disease in the flea-rat community in space and time. The sensitivity of the spread rate to the initial population size and spatial distribution is studied and the simulation results are compared with the theoretical ones.

Bennett et al. (2009) investigate a spatially explicit, individual-based, model of "disturbance activities," to explore effects of the spatial patterns of anthropogenic disturbance on the wildlife dynamics, for two case studies of avian species. They compare the influence of the pathways in proximity to and at the periphery of the nesting and foraging habitats of the yellow-headed blackbird and, in the second case study, the impact of unrestricted movement of recreationists on a breeding colony of Barbastelle bats.

Hynes et al. (2009) analyze the effects of a carbon equivalent tax on average family farm income, at the level of the individual farm and over the region. Simulating annealing techniques are employed to statistically match the Irish Census of Agriculture to a sample of representative farms and thus to generate a synthetic population of Irish farms. Their spatial micro-simulation indicates a heterogeneity in the farm population and demonstrates that (1) there would be significant regional variation in the burden of an agricultural tax, if based on a rate per unit of methane emissions; and (2) redistribution of the methane tax revenue, as an environmental subsidy, would encourage farmers to participate in the proposed agri-environmental scheme and improve the state of the low income farms.

Acknowledgment

This paper was supported by THE ISRAEL SCIENCE FOUNDATION (grant No. 692/06).

References

Al-Ahmadi, K., See, L., Heppenstall, A., Hogg, J., 2009. Calibration of a fuzzy cellular automata model of urban dynamics in Saudi Arabia. Ecological Complexity 6, 80–101.

Benenson, I., 2007. Editorial: WARNING! The scale of land-use CA is changing! Computers, Environment and Urban Systems 31, 107–113.

Bennett, V.J., Beard, M., Zollner, P.A., Fernández-Juricic, E., Westphal, L., LeBlanc, C.L., 2009. Understanding wildlife responses to human disturbance through simulation modelling: a management tool. Ecological Complexity 6, 113–134.

Hynes, S., Morrissey, K., O'Donoghue, C., Clarke, G., 2009. A spatial micro-simulation analysis of methane emissions from Irish agriculture. Ecological Complexity 6, 135–146.

Laperrière, V., Badariotti, D., Banos, A., Müller, J.P., 2009. Structural validation of an individual-based model for plague epidemics simulation. Ecological Complexity 6, 102–112.

Lee, D.B., 1973. A Requiem for large scale modeling. Journal Of The American Institute Of Planners 39 (3), 163–178.

Lee, D.B., 1994. Retrospective on large-scale urban models. Journal of the American Planning Association 60 (1), 35–40.